

Mintrop, Heinrich; Sundermann, Gail

Zentrale Steuerung von Schulentwicklung mit Quoten und Sanktionen: Das US-amerikanische Großexperiment in Schulen für benachteiligte Bildungsschichten

Die Deutsche Schule 104 (2012) 1, S. 8-30



Quellenangabe/ Reference:

Mintrop, Heinrich; Sundermann, Gail: Zentrale Steuerung von Schulentwicklung mit Quoten und Sanktionen: Das US-amerikanische Großexperiment in Schulen für benachteiligte Bildungsschichten - In: Die Deutsche Schule 104 (2012) 1, S. 8-30 - URN: urn:nbn:de:0111-pedocs-257189 - DOI: 10.25656/01:25718

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-257189>

<https://doi.org/10.25656/01:25718>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Heinrich Mintrop/Gail Sunderman

Zentrale Steuerung von Schulentwicklung mit Quoten und Sanktionen: Das US-amerikanische Großexperiment in Schulen für benachteiligte Bildungsschichten

Zusammenfassung

Seit mehr als zwei Jahrzehnten erproben die Regierungen der Vereinigten Staaten von Amerika Systeme zur Sicherstellung und Kontrolle von Schulqualität. Um die Schulentwicklung zentralistisch zu steuern und voranzubringen, setzt man vor allem auf die Wirkung quantitativer Wachstumsziele, verbunden mit konsequenten Sanktionierungen. Die Design-Merkmale dieses Systems haben tatsächlich Auswirkungen auf Schulen und Verwaltungen gehabt. Allerdings haben sie nicht zu den beabsichtigten Resultaten geführt, sondern stattdessen unerwartete organisationale Dysfunktionalitäten erzeugt und zu einer unerwünschten Lernkultur geführt.

Schlüsselwörter: Schulqualität, Leistungsanreize, Lehrermotivation, Schulentwicklung, Rechenschaftssystem, Zentralisierung

Centralized Management of School Improvement with Quota and Sanctions: The US-American Experiment with High-Stakes Accountability in Schools Serving Disadvantaged Populations

Abstract

In the last two decades or so, governments in the United States have experimented with systems of school quality assurance and control that heavily rely on the power of quantitative growth targets combined with stringent sanctions in guiding and incentivizing school development efforts from the center. The design features of these systems have indeed triggered pro-active responses on the part of schools and local administrative units, as intended. Yet, they have been unable to achieve their intended goals. Instead, they produced unexpected organizational dysfunktionalities and resulted in an undesirable learning culture.

Key words: accountability systems, performance incentives, school quality, work motivation, school improvement, centralization

In vielen Ländern durchleben wir derzeit eine Phase durchgreifender Zentralisierung in der Bildungspolitik, in der höhere Ebenen der Bildungsadministration zunehmend die Initiative bei der Bewältigung von Problemen der Schulentwicklung übernehmen und das traditionelle Bildungsestablishment – Lehrerschaft, Standesvertretungen, Universitäten – zu Nebenakteuren in ihren ureigensten Aufgabenkreisen machen. In den USA hat diese Entwicklung schon in den achtziger Jahren begonnen und in den letzten zwanzig Jahren enorm an Intensität gewonnen. Den Höhepunkt dieser Zentralisierung markiert die Verabschiedung eines Bundesgesetzes im amerikanischen Kongress im Jahre 2001 mit dem vielversprechenden Namen „No Child Left Behind“ (NCLB). Vor zehn Jahren von einer großen Koalition von Demokraten und Republikanern, Wirtschaftsverbänden und Bürgerrechtsorganisationen mit hohen Erwartungen aus der Taufe gehoben (vgl. Debray 2006), sind die wesentlichen Instrumente des NCLB-Bildungsreformwerks offenbar ausgereizt. Inzwischen sind Hoffnung und Überschwang in Ablehnung umgeschlagen.¹

In diesem Beitrag behandeln wir das US-amerikanische System als eine spezielle Variante zentraler Steuerung von Schulentwicklung und erklären sein besonderes Design vor dem Hintergrund des nationalen Kontextes. Wir untersuchen seine Auswirkungen auf Verwaltungen und Schulen über einen Zeitraum von zehn Jahren und diskutieren im Anschluss daran, welche Lehren für Qualitätssicherungssysteme aus den Erfahrungen gezogen werden können.

1. Der nationale Kontext

Dass es in den USA überhaupt zu einem bedeutenden System der zentralen Steuerung von Schulqualität kommen konnte, ist keineswegs eine Selbstverständlichkeit. In einem Forschungsbeitrag im Jahre 1990, der verschiedene Versuche der Einflussnahme von Politik auf das Schulleben resümierte, rief die bekannte Schulforscherin Darling-Hammond (1990) die „Macht der Basis über die (hohe) Politik“ aus. Damit brachte sie den Konsens der amerikanischen Schulforschung zum Ausdruck, dass stetig immer neue Initiativen und Maßnahmen von Seiten der Politik am Klassenraum abprallten. Großangelegte Implementationsforschungen zu zentralen Initiativen zur Unterrichtsreform konnten nur geringe Umsetzungseffekte erkennen und kamen zu dem Schluss, dass die Politik nicht vorgeben könne, was für Schulen relevant sei (vgl. McLaughlin 1990). Die Begründung für diese Machtlosigkeit zentraler Steuerung lieferte die Organisationssoziologie mit ihrer Analyse einer Struktur der losen Kopplung, die es dem System ermögliche, die Praxis vor Ort, also den Unterricht,

¹ Vgl. URL: <http://www.nytimes.com/2011/10/12/education/>.

von der administrativen und formalen Struktur der Organisation zu lösen. Die politische Zersplitterung des Systems tue ihr Übriges. Da Bildungspolitik in den USA eine Angelegenheit der Bundesstaaten sei, aber anders als in Deutschland traditionell an lokale Schulbehörden mit ihren eigenen gewählten Schulausschüssen delegiert werde, ergebe sich ein Markt der Ideen, auf dem Reformen sich eher wie ständig wechselnde oberflächliche Moden verbreiteten (vgl. Wirt/Kirst 1997; Cuban 1990; Cohen/Spillane 1992). Je mehr sich das System auf der Oberfläche verändere, desto mehr bleibe es sich im Kerne treu.

Trotzdem waren Strukturelemente einer administrativen Kontrolle vorhanden (vgl. Weiler 1990), die der Bildungshistoriker Tyack (1974) in seiner einflussreichen Studie der „*Administrative Progressives*“ des späten neunzehnten und frühen zwanzigsten Jahrhunderts herausgearbeitet hat: die Messung von Schulleistung mit Hilfe von quantitativen standardisierten Tests, die Dominanz spezialisierter (im 19. Jahrhundert männlicher) Administratoren über mehrheitlich weibliche Lehrkräfte und die herausragende Rolle empirischer Befunde zur Begründung von Reformen, die von externen wissenschaftlichen Instanzen erhoben werden. Im Zusammenspiel dieser Elemente ist Schulreform im US-Schulsystem häufig mit Bemühungen verbunden, das *scientific management* des Systems im Hinblick auf eine höhere Effizienz zu verbessern. Doch der „Kult der Effizienz“ (Callahan 1962), die Vorstellung, dass die Feinstruktur des Unterrichts generell administrativ regulierbar sei oder sein solle, ist nicht auf Manager beschränkt, die auf Effizienz bedacht sind. Progressive Pädagogen und Pädagoginnen, Wissenschaftler und Wissenschaftlerinnen, die „*Best Practice*“-Unterrichtsprogramme entwickeln (vgl. z.B. Slavin 2005), oder auch politische Akteure, die für mehr Chancengleichheit eintreten (vgl. The Education Trust 2004), setzen auf sogenannte *instructional policy* und externes *instructional management* für die Lösung unterrichtlicher Probleme.

Administrative Regulierung sollte aber nicht mit zentraler Steuerung gleichgesetzt werden. Traditionell verfügen die Bildungsministerien in den USA auf der Ebene der Bundesstaaten nur über kleine administrative Apparate, die kaum Kapazitäten jenseits der Überprüfung von formalen Regeln, Gesetzen und lokalen Finanzbudgets zur Verfügung haben. Im Staat Kalifornien z.B. laufen die Apparate der kommunalen Schulbehörden in den großen Städten bei weitem dem Bildungsministerium den Rang ab (vgl. Timar 1997). So waren die ersten Ansätze verstärkter staatlicher Steuerung in den 80er-Jahren denn auch eher traditionell (vgl. Timar/Kirp 1988; Weiler 1990): staatliche Standards, neue Lehrpläne, Stärkung der Hauptfächer usw.

Die zweite Welle der Zentralisierung in den 90er-Jahren brachte etwas völlig Neues. Sie hatte drei Veränderungen, im Finanzwesen, der Arbeitsorganisation und der administrativen Kontrollkapazität, zur Voraussetzung:

- Fiskalisch ging in vielen Bundesstaaten, häufig als Folge von gerichtlich erzwungenen Angleichungen zwischen armen und reichen Kommunen, die Zuständigkeit für die Finanzierung des Schulwesens von den Kommunen auf die Bundesstaaten über. Die Folge war, dass zentrale Staatsregierungen neue fiskale Autorität erhielten (vgl. Verstegen/Knoepfel 1998).
- Beginnend in den 80er-Jahren wich die Arbeitsorganisation des „Organisationsmenschen“ (vgl. Whyte 1956) der des „Hochleistungsmenschen“ (vgl. Sennett 2006). Funktionierte der Organisationsmensch in stabilen öffentlichen und privatwirtschaftlichen Bürokratien, in denen Loyalität der Beschäftigten und das Prinzip der Seniorität vorherrschten, geht es in der heutigen Organisation von Arbeit um höhere Einsätze (*stakes*). Hochleistung wird belohnt, schwache Leistung führt zur Kündigung. Diese auf Hochleistung ausgerichtete Arbeitsorganisation ist zur vorherrschenden Leitfigur in der amerikanischen Privatwirtschaft geworden und wurde in den 90er-Jahren auf das öffentliche Schulsystem der USA übertragen.
- Eine dritte Vorbedingung sind neue Kontrollmechanismen vor allem für ehemals nur schwer kontrollierbare komplexe Arbeiten (vgl. Sennett 2006; Oldham/Hackman 2010). Mit Hilfe von neuen computergestützten Techniken der Datenerhebung und -speicherung ist es nun einer zentralen Planungsbehörde, sei sie öffentlich oder privat, möglich, klare Leistungszielvorgaben zu formulieren (*goal setting*) und mit relativ geringem Aufwand ihre Einhaltung zu überprüfen (*goal monitoring*). Die neuen Techniken erlauben es und machen es erforderlich, Ziele mit begrenzten quantitativen Leistungsindikatoren zu erfassen, zentrale Leistungserwartungen für eine große Anzahl von kleinen Leistungs- oder Arbeitseinheiten zu überprüfen und jeweils zielgenaue Sanktionen für leistungsschwache Einheiten anzuordnen. Was ehemals mittlere Vermittlungsebenen besorgten, liegt nun in der Hand des zentralen Leistungsmanagements. Im Bereich der Politik geben diese Verwaltungsstrukturen schmalen, zentral positionierten Eliten, egal ob eher effizienzorientiert oder sozial-reformerisch, neue machtvollere Steuerungsmöglichkeiten an die Hand.

Neue staatliche fiskalische Autorität, neue Arbeitsorganisation und neue Kontrollmöglichkeiten machten es nun möglich, Schulentwicklung über ein zentrales Rechenschaftssystem zu steuern.

2. Qualitätssicherung als Rechenschaftssystem für Leistung

Die gängigen Rechenschaftssysteme für Schulen in den USA bestehen von ihrer Konstruktion her aus: (a) kompetenz-, inhalts- und fertigungsorientierten Standards, (b) standardisierten Tests, die die Basis für Schul-Leistungsindikatoren sind, (c) quantitativen Zielvorgaben und Leistungsquoten (z.B. im Jahre X haben Y Prozent der Schüler und Schülerinnen Minimalstandards zu erfüllen), (d) Sanktionen, mit denen sich die höheren Politik- und Systemebenen bei den jeweils niedrigeren Gehör

verschaffen und (e) Unterstützungs- und Interventionssystemen, die neue Mittel und Ideen an die bedürftigsten Schulen verteilen sollen (vgl. Mintrop/Sunderman 2009; Fuhrman/O'Day 1996; O'Day 2002).

Das NCLB-Regelwerk fordert, dass Schulen in einem festgesetzten Zeitraum – es war der Zeitraum von 2001 bis 2014 geplant – alle Schüler und Schülerinnen auf das Leistungsniveau „*proficient*“ (fortgeschrittene Kenntnisse) zu bringen hatten. Aus der Differenz zwischen diesem Ziel und der tatsächlichen durchschnittlichen Leistung einer Schule, gemessen in jährlichen staatlichen Tests, ergibt sich die ‚angemessene jährliche Zuwachsrate‘ (*adequate yearly progress, AYP*). Für jede statistisch relevante demografische Untergruppe, differenziert nach Einkommen, Ethnie, Muttersprache und Förderungsbedarf, wird eine schulspezifische, jährlich zu erreichende Testquote kalkuliert. Dies soll die Schulen dazu zwingen, schwarze, hispanische, einkommensschwache und lernbehinderte Kinder genauso zu fördern wie weiße oder Kinder aus der Mittelschicht. Die Organisationseinheiten, die zur Rechenschaft gezogen werden, sind die Einzelschule und die kommunale Schulbehörde, nicht individuelle Lehrkräfte und auch nicht Eltern oder Schüler und Schülerinnen.

Die inhaltliche Substanz der Standards, das Anforderungsprofil der Tests, die Definition von „*proficient*“, die Bereitstellung von Inputs, die Einbindung von Schülern und Schülerinnen in das Sanktionssystem (d.h. wie Benotung und Versetzung mit den bundesstaatlichen Tests verknüpft werden sollten) und die spezielle Art und Weise der Unterstützung und Intervention in leistungsschwachen Schulen – alles das war der jeweiligen *einzelstaatlichen* Regulierung anheimgestellt.² Auf nationaler Ebene dagegen sind Sanktionsstufen vorgeschrieben, die eine Schule oder Schulbehörde durchlaufen muss, wenn sie die Zielvorgaben über einen Zeitraum von fünf Jahren nicht erfüllt. Auf der ersten Stufe steht die Veröffentlichung der schwachen Schulleistung, die die Schule oder Schulbehörde dem Druck ihres Marktes aussetzt. Auf der zweiten Stufe verliert die Schule oder Schulbehörde zunehmend ihre Autonomie und muss sich externen Interventionen und privaten Anbietern bestimmter Funktionen öffnen und dies auch bezahlen. Auf der dritten Stufe wird die Schule geschlossen, oder die kommunale Schulbehörde wird von der Staatsregierung übernommen.

Die Anhebung der Leistungsstandards eines großen nationalen Bildungssystems und insbesondere eine dramatische Verbesserung des Bildungsstandes von Gruppen mit niedrigem gesellschaftlichem Status sind enorm komplexe Angelegenheiten. Das gegenwärtig gültige amerikanische Rechenschaftssystem für Schulen hingegen besticht durch seine Simplizität. Im Kern besteht das System aus Quoten und Sanktionen. Und doch sind in Anbetracht der Bildungsmängel, mit denen zunehmend ausgegrenzte

2 Gemäß der föderalen Struktur in den USA sind die Staaten für die Bildung verantwortlich. Bundesgesetze, wie das NCLB-Gesetz, erreichen ihre Wirkung nur, wenn Staatsregierungen im Austausch gegen Bundesmittel in neue bundesweite Regeln und Kontrollen einwilligen.

Kinder aus den unteren Sozialschichten und den ethnischen Minoritäten zu kämpfen haben, Rechenschaftssysteme, die eine positive Bewegung in die Schulentwicklung bringen, grundsätzlich begrüßenswert. Entscheidend ist jedoch, dass sie, auch wenn sie nicht in allen Einzelheiten austariert sein sollten, den Bildungsauftrag der Schule im Wesen nicht verzerren. Von diesen Überlegungen lassen wir uns leiten, wenn wir im Folgenden die Auswirkungen des Systems auf die unteren Ebenen untersuchen.

3. Gemessene Leistungseffekte

Im Lichte wissenschaftlicher Studien ist es umstritten, ob die Rechenschaftssysteme die versprochene Verbesserung der messbaren Schülerleistung erbringen. Schon die Untersuchung der staatlichen Systeme der 90er-Jahre führte zu Kontroversen (vgl. Amrein/Berliner 2002; Carnoy/Loeb 2002; Haney 2000; Raymond/Hanushek 2003). Auch später ergibt sich kein klares Bild. Effektbefunde variieren von Studie zu Studie; und dort, wo Effekte gefunden werden, bleiben die kausalen Mechanismen zwischen Testergebnissen und Systemdesign im Dunkeln (vgl. Lee 2007; Nichols/Glass/Berliner 2006). Oftmals hängen gemessene Effekte von methodischen Entscheidungen der Forschenden ab (vgl. Braun 2004). Dabei geht es insbesondere um die Frage, wie mit Ausschluss von getesteten Schülern und Schülerinnen statistisch umgegangen werden soll. Eine Metaanalyse, gestützt auf eine Reihe großangelegter Studien mit nationalen Datensätzen, kommt zu dem Resultat, dass es einen mäßig positiven Effekt für die Accountability-Periode der Schulreform gegeben habe, aber keine signifikante Reduktion der Leistungsspanne zwischen oberen und unteren Statusgruppen (vgl. Lee 2008).

Gemessene Effekte und deren Interpretation hängen auch davon ab, welche Messeinheit zugrunde gelegt wird. Die meisten Rechenschaftssysteme der einzelnen Staaten erscheinen auf den ersten Blick erfolgreich zu sein. Studien, die die staatlichen Tests zur Grundlage nehmen, finden fast durchweg einen kontinuierlichen, oftmals steilen Anstieg der durchschnittlichen Testleistungen (vgl. Center for Education Policy 2008b; Neal/Schanzenbach 2008; The Education Trust 2004, 2005).

Ein bei weitem weniger positives Bild ergibt sich, wenn man die Ergebnisse des *National Assessment of Educational Progress* (NAEP) zugrunde legt, die in regelmäßigen Abständen als Zufallsstichprobe in allen Staaten erhoben werden. An das NAEP sind keine Sanktionen für Schulen oder Behörden geknüpft; NAEP ist ‚low stakes‘. Ein Vergleich von Testergebnissen der einzelnen Bundesstaaten und NAEP kann klären, wie realistisch die gemessenen Lernzuwächse wirklich sind. Um Verzerrungen (z.B. curriculare Verengungen, übertriebene Wiederholungen von Testitems) als Ursache für Testresultate auszuschließen, sollten *High-Stakes*- und *Low-Stakes*-Tests ein konsistentes Bild vermitteln (vgl. Koretz 2008; Lee 2007). Wenn man NAEP-Resultate zu-

grunde legt, fällt der Zuwachs in den Testergebnissen jedoch sehr viel geringer aus (vgl. Fuller u.a. 2006; Fuller u.a. 2007; Lee 2006, 2007). Generell sind die Ergebnisse für die Grundschulen besser als für weiterführende Schulen und in Mathematik etwas besser als im Bereich der Lesefähigkeiten. Entwickelte Steuersysteme, in denen Standards, Tests und Unterrichtsmaterialien aufeinander abgestimmt sind, scheinen positive Auswirkungen zu haben (vgl. Carnoy/Loeb 2002). Resümierend muss jedoch festgestellt werden, dass die großen Unterschiede zwischen den ethnischen und sozioökonomischen Statusgruppen weiter bestehen (vgl. Lee 2007, 2008). Zehn Jahre nach der Einführung des NCLB-Regelwerks ist klar geworden, dass das quantitative Wachstumsziel des Systems, dass die gesamte Schülerschaft das Leistungsniveau „*proficient*“ erreicht, in allen Bundesstaaten weit verfehlt worden ist.

4. Anreizfunktion und Arbeitsmotivation

In Anbetracht der weiten Diskrepanzen zwischen NAEP (Systemüberprüfung) und den jeweiligen bundesstaatlichen Tests (Rechenschaftslegung) ist es nicht geklärt, was die letzteren eigentlich messen. So viel scheint klar: Die staatlichen Rechenschaftssysteme, die mit Anreizen, Druck und Sanktionen operieren, schaffen es, die Aufmerksamkeit der Schulen auf die Tests zu richten und sie zu einer sorgfältigen Ausrichtung (*alignment*) auf die Standards und Testinhalte zu motivieren (vgl. Koretz 2008).

Die Übernahme der Verantwortung für die Leistungsentwicklung der Schülerinnen und Schüler wird von der Lehrerschaft als Idee im Allgemeinen akzeptiert, aber die konkreten Überprüfungssysteme werden eher skeptisch betrachtet (vgl. Conley/Goldman 2000; Darling-Hammond 2004; Firestone/Monfils/Schorr 2004; Hill/Lake/Celio 2002; Koretz u.a. 1996; Leithwood/Steinbach/Jantzi 2002; Loeb/Knapp/Elfers 2008; Mintrop/Trujillo 2007; Spillane u.a. 2002). Lehrkräfte halten eine begrenzte Orientierung des Unterrichts an den Tests (*teaching to the test*) für akzeptabel, lehnen aber den starken Druck und die damit verbundene Vernachlässigung von weniger leicht messbaren Lernzielen ab (vgl. Abrams/Pedulla/Madaus 2003; McNeil 2000; Valenzuela 2005).³ Jedoch greifen die Überprüfungssysteme insofern, als Schulen sie zur Orientierung für die Unterrichtsgestaltung heranziehen und energisch versuchen, den drohenden Sanktionen zu entkommen (vgl. Finnigan/Gross 2007; Mintrop 2004; Kelley/Protsik 1997).

Oftmals fühlen sich die engagiertesten Lehrkräfte im Kollegium, die am wenigsten Grund hätten, sich als Versager zu fühlen, am meisten dafür verantwortlich, die Leistungsbilanz der Schule zu verbessern. Doch wenn die erhofften Erfolge ausblei-

3 In der Tat konnte in einer Fallstudie kalifornischer Schulen ein enger Zusammenhang zwischen Testergebnissen, dem Leistungsrang einer Schule im System und der Qualität von Unterricht nicht nachgewiesen werden (vgl. Mintrop/Trujillo 2007).

ben oder sich nicht verstetigen, nachdem kurzfristige Lösungen erprobt worden sind, wie dies oft in leistungsschwachen Schulen mit geringer Entwicklungskapazität der Fall ist, dann breitet sich Demoralisierung aus (vgl. Finnigan/Gross 2007; Mintrop 2004). Einige Studien berichten von Schulen und Schulbehörden, die sich mit dem Rückenwind des Systems rapide aufwärts bewegen konnten (vgl. Skrla/Scheurich 2004; Hill/Lake/Celio 2002). Dennoch scheinen diese positiven Fälle nicht die Regel zu sein (vgl. Sunderman u.a. 2004; Abrams/Pedulla/Madaus 2003; Louis/Febey/Schroeder 2005). Vielmehr macht sich ein überwältigendes Klima der Negativität im unteren Drittel des Schulleistungsspektrums breit.

5. Praktikabilität und Lernkultur

Auf der Systemebene ist der Erfolg des Anreizmechanismus vom Zusammenspiel zwischen der Realitätsangemessenheit der Zielvorgaben, den vorhandenen Kapazitäten oder Kapazitätsreserven der unteren Ebenen, dem erzeugten Druck und der internen Wertschätzung der Leistungsforderungen oder Belohnungen abhängig (vgl. Mintrop 2004; Kelly/Protsik 1997). Das Funktionieren – und politische Überleben – der Rechenschaftssysteme in ihrer momentanen Konstruktion ist unmittelbar von der Höhe der kognitiven Anforderungen in den staatlichen Tests abhängig. Diese These kann mit Michael Fullans Konzept der zwei Zielhorizonte (vgl. Fullan 2003), die er aus der Evaluation der englischen Schulreform entwickelte, plausibilisiert werden. Innerhalb des ersten Horizontes konzentrieren sich die Rechenschaftssysteme auf Basiswissen und -fertigkeiten. Davon ausgehend erfolgen eine strikte Ausrichtung (*alignment*) der Curricula einer Schule an den staatlichen Tests, präskriptive Unterrichtsprogramme und das Festzurren von Organisationsabläufen. Außerdem kommen alle möglichen kurzfristigen Strategien, wie die besondere Aufmerksamkeit auf Schüler und Schülerinnen an den statistischen Schnittpunkten, Training von Testaufgaben, das enge Wiederholen von einzelnen Standards, die emotionale Mobilisierung der Schüler und Schülerinnen kurz vor dem Test usw., zur Anwendung. Diese scheinen, zumindest am Anfang, einen positiven Einfluss auf die Testresultate zu haben. Aber die Zuwächse scheinen nicht nachhaltig zu sein (vgl. Fullan 2003). Innerhalb eines erweiterten Zielhorizontes nehmen die Systeme kognitive Ziele jenseits von abrufbarem und wiederholbarem Wissen ins Visier, wie z.B. Begründen und komplexes Problemlösen.

Staatliche Rechenschaftssysteme, deren Tests auf relativ niedrigere kognitive Anforderungen abheben, wie z.B. im Staat Texas, der mehr als andere Staaten dem NCLB-Regelwerk Pate stand, haben innerhalb des gegenwärtigen NCLB-Regelwerks „besser“ funktioniert. Da sie auf existierende Kompetenzen von Lehrkräften und Schülern und Schülerinnen im unteren Bereich zugeschnitten sind, produzieren diese Systeme weniger Schulen, die als mangelhaft ausgewiesen sind, und brauchen gerin-

gere, weniger kostspielige Maßnahmen im Bereich der Schulentwicklung oder Lehrerfortbildung. Druck mag in diesem Fall als Steuerungsmechanismus hinreichend sein, um die Schulen auf die relativ moderaten staatlichen Ziele zu verpflichten. Im Kontrast dazu findet man staatliche Systeme, z.B. in Kalifornien, die auf höhere kognitive Leistungen abzielen und dabei völlig impraktikabel werden (vgl. Mintrop 2008; Mintrop/Trujillo 2005).

Wenn nun immer mehr Bundesstaaten es dem Staat Kalifornien gleichtäten und ihre Anforderungen für das Leistungsniveau „*Proficiency*“ den NAEP-Niveaus des Bundes angleichen, wie es auch von vielen Reformern und Reformerrinnen gefordert wird, dann hätte das den Nachweis einer enormen Anzahl von mangelhaften Schulen zur Folge, für die gewaltige Interventionskapazitäten bereitgestellt werden müssten. In der Tat zeigt Kalifornien, ein Staat mit hoher Einwanderungsquote und einer Mehrheit von nicht-weißen Schülern und Schülerinnen, wohin die Entwicklung laufen könnte. Im Jahre 2008 waren dort schon 48 Prozent aller Schulen und 61 Prozent aller kommunalen Schulbehörden von Bundessanktionen betroffen (vgl. Asimov 2008). Die Alternative zu diesem Szenario ist, die Rechenschaftssysteme im sicheren Hafen des unteren Zielhorizontes zu belassen, statt totale Dysfunktionalität zu riskieren. In der Tat haben Paterson und Hess (2008) eine Tendenz nach unten – „*race to the bottom*“ – diagnostiziert, der zufolge Staaten den starren Wachstumsanforderungen des Bundes zunehmend mit einfacheren Leistungskriterien für „*Proficiency*“ begegnen.

Ziele im Bereich des unteren Zielhorizontes sind eine lohnende Herausforderung für viele Schulen. Doch sind viele Schul- und Unterrichtsentwicklungsstrategien, die auf diesen Horizont ausgerichtet sind, wenig geeignet, Ziele im höheren Leistungsbereich zu erlangen. Kognitive Leistungsanforderungen im Basisbereich, gekoppelt mit hohem Druck, haben kommunale Schulbehörden und Schulleitungen, vor allem im eher leistungsschwachen Spektrum, dazu veranlasst, präskriptive, engmaschige Unterrichtsprogramme zu verordnen, die auf die staatlichen Tests zugeschnitten sind und die, tayloristisch, mit einer buchstäblichen Bindung an das Lehrermanual, einer so genannten „*implementation fidelity*“, zu unterrichten sind (vgl. Protheroe 2008). Das Schulwissen hält sich eng an die Testvorgaben, wird fragmentiert und stark lehrerzentriert vermittelt (vgl. Au 2007; Herman 2004). Das politische Ziel, allen Schülern und Schülerinnen die Möglichkeit zu geben, am „Lernen des 21. Jahrhunderts“ (vgl. <http://www.p21.org>; Zugriffsdatum: 13.02.2012) teilzunehmen, rückt so in weite Ferne.

6. Simplizität und Identifizierung von schwachen Schulen

Das NCLB-Regelwerk, wie übrigens auch die meisten Rechenschaftssysteme der Einzelstaaten, besticht durch seine Simplizität, die von der zentralen Administration wenig Datenkapazität verlangt, den lokalen Ebenen wenig Spielraum für Interpretationen und Zweideutigkeiten lässt und der Öffentlichkeit eine eingängige Metrik an die Hand gibt. Gemäß dem NCLB-Gesetz werden Entwicklungsfortschritte von Schulen in simplen Kompetenzstufen gemessen: „(far) below basic“, „basic“ und „proficient“. Bis dato sind aber Sanktionen nur von einem einzigen Kriterium abhängig, nämlich ob oder bis zu welchem Grade die Schulen einen jährlich zunehmenden Anteil von Schülern und Schülerinnen zur Kompetenzstufe „proficient“ geführt haben, und zwar geltend für alle ethnischen und sozio-ökonomischen Schüleruntergruppen. Dies hat zur Folge, dass Bewegungen im Bereich der Kompetenzstufe „basic“, vor allem nahe des „Proficiency“-Schnittpunkts, registriert werden, während das System positive Entwicklungen, z.B. im Bereich „far below basic“, wo viele leistungsschwache Schulen aber operieren, ignoriert (vgl. Neal/Schanzenbach 2008). Tatsächlich sind die durchschnittlichen Lernzuwächse in Schulen, die für Sanktionen infrage kommen, oftmals ähnlich denen von Schulen, die ihr AYP erreicht haben, wie eine Studie von Daten des Staates Virginia belegt (vgl. Kim/Sunderman 2005). Diese Ungenauigkeit wäre vielleicht tolerierbar, wenn AYP eines von mehreren Kriterien der Schulleistungsmessung wäre. Doch dies ist nicht der Fall. AYP als der einzige autoritative Indikator, der den Sanktionen zugrunde gelegt wird, schafft Realitäten und hat weitreichende Konsequenzen für Schulen und Schulbehörden.

In Anbetracht dessen, dass außerschulische Faktoren Schülerleistungen stark beeinflussen (vgl. Scheerens/Bosker 1997), können Schulen ihre Chancen im System beträchtlich steigern, indem sie die Zusammensetzung der Schülerschaft, die zum Test antritt, manipulieren. Eine Studie, basierend auf nationalen Daten (vgl. Balfanz/Legters 2008), hat Schulen untersucht, in denen sich die Jahrgangskohorte zwischen der 9. und 12. Klasse um mindestens 40 Prozent verringert. Nach den Kriterien des NCLB-Regelwerks wurden 41 Prozent dieser Schulen als akzeptabel eingestuft, da sie ihre AYP-Vorgaben erfüllten (vgl. Balfanz/Legters 2004, 2008), obwohl eine große Anzahl der Schüler und Schülerinnen die Schule ohne Abschluss verlässt. Ausschluss-Strategien, wie z.B. zunehmendes Sitzenbleiben-Lassen vor wichtigen Jahrestests, Abschiebung von schwachen Schülern und Schülerinnen in nicht-rechenschaftspflichtige Sonderschul-Klassifikationen oder *Push-outs*, d.h. eine Tolerierung von oder Ermutigung zu Schulmüdigkeit bei schwachen Schülern und Schülerinnen (vgl. Haney 2008; Booher-Jennings 2005; Anagnostopoulos 2003), sind als Schnelllösungen in einem System naheliegend, welches Erfolg oder Misserfolg, Belohnung oder Strafe lediglich an dem Anteil der Schülerschaft bemisst, der es bis zur „Proficiency“-Kompetenzstufe geschafft hat.

7. Standardisierung und spezielle Lernbedürfnisse

Das Rechenschaftssystem verpflichtet alle Schulen, die Gelder vom Bund annehmen, Testergebnisse nach Schüleruntergruppen zu disaggregieren. Dies hat die Transparenz vor allem hinsichtlich der Leistungen von benachteiligten Schülern und Schülerinnen erhöht und Druck auf die Schulen erzeugt, Leistungsschwächen von Schülern und Schülerinnen, die ethnischen Minoritäten angehören, aus einkommensschwachen Familien stammen, Lernbehinderungen aufweisen oder eine nicht englische Herkunftssprache haben, ernst zu nehmen. Aber das Vorgehen hat auch neue, unerwartete Nebeneffekte erzeugt. Leistungsschwache Schulen haben häufig eine hohe Anzahl von Untergruppen, was zu vielen (maximal bis zu 37) verschiedenen Leistungsvorgaben (vgl. Marion u.a. 2002) führen kann. Je höher die demografische Diversität ist, desto höher ist die statistische Wahrscheinlichkeit, dass eine Vorgabe nicht erfüllt werden kann und Sanktionen folgen (vgl. Kim/Sunderman 2005; Linn 2005; Novak 2003).

Es hat sich herausgestellt, dass es den Schulen besondere Schwierigkeiten bereitet, Zielvorgaben für Kinder in zwei Kategorien zu erreichen, nämlich für solche mit nicht-englischer Herkunftssprache, „*limited English proficiency*“ (LEP), und solche mit Lernbehinderungen, „*learning disabilities*“ (LD). Die Folge ist, dass viele solide arbeitende Schulen unter das Sanktionsregime fallen, weil sie ihre LEP- oder LD-Untergruppen nicht nahe genug an die „*Proficiency*“-Marke herangebracht haben. Die Forschung über Rechenschaftssysteme im Bereich LEP ist relativ solide. Verschiedene Implementierungsprobleme haben sich herausgeschält, vor allem der ständige Austausch von leistungsstärkeren Abgängen mit Neuzugängen mit großen sprachlichen Defiziten und die mangelnde Diagnostik der standardisierten staatlichen Tests für Zweitsprachler und -sprachlerinnen (vgl. Abedi 2004; Batt/Kim/Sunderman 2005; Coltrane 2002; Kieffer/Lesaux/Snow 2008). Ein ähnliches Muster ergibt sich für die „*Learning-Disabilities*“-Untergruppe (vgl. Batt/Kim/Sunderman 2005; Sunderman/Kim/Orfield 2005).

Es sollte nicht überraschen, dass ein System, welches Schulleistung mit hoch standardisierten und simplen Einheiten bemisst und wenig Abweichungen von einheitlichen Zielvorgaben erlaubt, bei der Behandlung von Schülern und Schülerinnen mit besonderem Lernbedarf unter Stress gerät. Aber die Lösung dieses Problems ist nicht einfach. In der Logik von Druck und Sanktionen werden Gruppen, die von der Rechenschaftslegung ausgenommen werden, etwa weil sie nur schwerlich standardisiert werden können, leicht ignoriert. Es ist nun einmal im Selbstinteresse der Schulen und der Schulbehörden, sich in ihrer Arbeit an den Forderungen mit dem größten Rechenschaftsdruck auszurichten.

8. Radikale Interventionen und Schulentwicklung

Staatsregierungen und lokale Schulbehörden vermeiden die radikalen Maßnahmen zur Umstrukturierung, die von der Bundesregierung im NCLB-Gesetz vorgeschlagen werden, wie z.B. die Schließung und Neueröffnung einer Schule mit neuem Kollegium (*reconstitution*) oder die Umwandlung einer öffentlichen in eine privat geführte Schule (vgl. Center for Education Policy 2007, 2008a; GAO 2007), zugunsten von weicheren, eher traditionellen Schulentwicklungsmaßnahmen (vgl. Center for Education Policy 2007). In der Tat haben Studien gezeigt, dass die radikaleren Methoden nicht immer die besseren sind. Bei der Rekonstituierung einer Schule mit neuem Kollegium, dies dokumentieren Studien aus Maryland (vgl. Malen u.a. 2002), Chicago (vgl. Hess 2003) und New York (vgl. Brady 2003; NYSED 2003), profitieren in einigen Fällen die Schulen vom Neuanfang, aber in vielen anderen Fällen werden die Schulen nicht besser; einige verschlechtern sich aufgrund von zusätzlichem Chaos. Die Übernahme von Schulen in private Trägerschaft (*educational management organizations*), die häufig eine größere Anzahl von leistungsschwachen Schulen managen, hat manchmal positive, manchmal negative Auswirkungen, wie Studien aus Baltimore und Philadelphia zeigen (vgl. Blanc 2003; Bracey 2002; Molnar 2005; Travers 2003a; Useem 2005; Saltman 2005). Trotz zunehmender Forschungen zu den so genannten *Charter Schools* – Schulen mit halböffentlichem Status, die oftmals von Eltern, Lehrkräften, Nachbarschaftsorganisationen, aber auch von den großen Managementorganisationen betrieben werden – gibt es wenig Anhaltspunkte, dass die Umwandlung von öffentlichen Schulen in halböffentliche *Charter Schools* eine hilfreiche Strategie ist (vgl. The Brown Center 2003; Gill u.a. 2007). Es ist zusammenfassend festzustellen, dass keine der Umstrukturierungsstrategien, mit denen bisher in den USA experimentiert wurde, sich als generell wirksam für Schulentwicklung erwiesen hat oder robust genug ist, sich gegen lokale Bedingungen durchzusetzen (vgl. auch Mathis 2009).

Die starre Abstufung von Sanktionsstufen hat für immer mehr Schulen unproduktive Turbulenzen produziert, anstatt eine sich selbst verstärkende, geordnete Entwicklungsdynamik hervorzurufen. Schulen, die über Jahre leistungsschwach gewesen sind, werden nun nicht mit den Sanktionen sozusagen „aus dem Schulentwicklungs-Dornröschenschlaf geweckt“. Ganz im Gegenteil, diese Schulen sind in ihrer Mehrzahl nicht stabil agierende Organisationen, sondern befinden sich in ständigem ungeplantem Auf- oder Abbruch. Lehrkräfte und Schulleitungen kommen und gehen; externe Berater und Beraterinnen haben sich an ihnen zu Hauf versucht, und den Bundessanktionen sind oftmals bereits intensive lokale Interventionen vorausgegangen (vgl. Mintrop 2004; Neild/Spiridakis 2002; Neild u.a. 2003). Die Stadt Philadelphia liefert ein gutes Beispiel. Hier wurden alle leistungsschwachen Schulen schon vor dem Inkrafttreten der Bundessanktionen radikal umgestaltet (vgl. Travers 2003b), ohne dass damit das Problem mangelnder Leistung beseitigt worden wäre.

Diese Schulen durchlaufen die NCLB-Sanktionsstufen, haben aber bereits schon die schärfsten Sanktionsstufen, wie radikale Neugründung oder neue Trägerschaft, vorweggenommen. Dazu kommt die Instabilität von kommunalen Schulbehörden, die bei dem ständigen Druck wegen krisenhafter Leistungen mit vermehrtem Abgang von Verwaltungsleuten und einem häufigen Wechsel an der Spitze zu kämpfen haben (vgl. Buchanan 2006).

Schulen in den städtischen Zentren, wo die Armen und die Minoritäten in den USA oftmals konzentriert leben, beschäftigen in zunehmendem Maße Lehrkräfte und Schulleitungen mit einer Fluktuationsrate von bis zu zwei Dritteln über einen Zeitraum von fünf Jahren, da die Beschäftigten nur schwer über einen längeren Zeitraum gehalten werden können (vgl. Johnson 2007). Bei diesen Lehrerkohorten geht es also weniger darum, wie sie zur Rechenschaft gezogen werden, als wie sie zum Bleiben bewegt werden können. Keine der Sanktionen oder Interventionen, weder auf der nationalen Ebene noch auf der bundesstaatlichen Ebene, scheint auf dieses Problem zu reagieren. Im Kontext der Hochleistungs-Arbeitsorganisation ist hingegen das regelmäßige „*Hiring and Firing*“ eine integrale und gewollte Dynamik.

9. Reform-Infrastruktur und Unterstützung

Forschungen haben gezeigt, dass es Schulen und Behörden gibt, die durchaus die notwendigen Kapazitäten haben, ihr Leistungsdefizit zu kompensieren. Schulen mit genügender Kapazität können so auf Druck und Orientierung des Rechenschaftssystems positiv reagieren (vgl. Diamond/Spillane 2004; Elmore 2004; O’Day 2002; Sunderman 2001). Jedoch mangelt es in vielen leistungsschwachen Schulen an professionellem Wissen, Geldern und Selbstorganisation. Diese Schulen sind nur schwer in der Lage, die Signale des Systems positiv für die Schulentwicklung umzusetzen (vgl. Elmore 2004). Ständig verstärkter Druck führt hier zu mehr Rigidität, Fragmentierung und Verschlechterung der kollegialen Beziehungen, was gemeinsames Lernen und Erproben von neuen Ansätzen stark behindert (vgl. O’Day 2002; Mintrop 2004). Schulen mit mangelnden Kapazitäten sind dazu prädestiniert, mit kurzfristigen Strategien zu operieren, die kaum neue Kapazitäten erfordern, wie z.B. eine enge Ausrichtung auf die Tests (vgl. Sunderman u.a. 2004). Im Extremfall kann dies zu einem dezidiert anderen Curriculum führen, in welchem ein testzentrierter Förderunterricht das reguläre Curriculum an den Rand schiebt (vgl. Booher-Jennings 2005; Diamond/Spillane 2004; Sunderman 2001; Valenzuela 2005).

Im Regelwerk der amerikanischen Bundesregierung spielen die bundesstaatlichen Bildungsministerien bei der Implementierung von NCLB eine zentrale Rolle, doch sind wenig neue Gelder für die Bewältigung der erweiterten Aufgaben dieser Ministerien bewilligt worden. Die „*State Departments of Education*“, traditionell rela-

tiv schwach besetzte Organisationen, haben weder Personal und Expertise noch etablierte Kommunikationskanäle, die für die Schulentwicklung benötigt werden (vgl. McDermott 2004; Sunderman/Orfield 2006), sodass sie in der Regel diese Aufgaben nach unten delegieren – an Behörden und Schulen also, die oftmals gerade deshalb den Sanktionen ausgesetzt sind, weil sie sich in der Vergangenheit aus eigener Kraft nicht verbessern konnten.

10. Zusammenfassung und Diskussion

Im Kern geht es bei Qualitätssicherung um guten Unterricht. Guter Unterricht basiert auf einer komplexen Interaktion von Wissen, Zielen, Diagnostik, Motivation, Anstrengung, Belohnung und Entscheidungsfindung, wobei Lehrkräfte lokale Einflüsse und Präferenzen mit allgemeineren professionellen Wissensbeständen und Normen vermitteln müssen. Gleichzeitig müssen lokale Kriterien übergeordneten gesellschaftlichen Leistungs-, Wissens- und Verhaltensstandards standhalten (vgl. auch O’Day 2002; Weiler 1990). Idealtypisch könnten Systeme der Qualitätssicherung lokal, professionell oder administrativ organisiert sein und zwischen diesen Ebenen vermitteln. Eine lokale Dynamik kommt zum Tragen, wenn lokale Eliten oder Eltern gesellschaftlich erforderliches Wissen überblicken können und selbst, oft in direktem persönlichen Kontakt, die Aufsicht über eine Lehrerschaft übernehmen, die im Konsens mit der lokalen Gemeinde agiert und eine eher geringe Kontrolle über ihre Arbeit hat. Professionelle Dynamiken der Rechenschaftslegung appellieren an etabliertes Wissen und verinnerlichte Erziehungs- und Leistungsnormen der Lehrerschaft, die in der Lehrerausbildung vermittelt werden und durch internen kollegialen Austausch in den Schulen oder mit Hilfe externer Experten und Expertinnen wach gehalten und verstärkt werden. Administrative Mechanismen beruhen auf externen Leistungskriterien, Regeln, Sanktionen und Ressourcen für Lehrende und Lernende, die in wechselseitiger Interaktion zwischen der Zentrale und den lokalen, nachgeordneten Ebenen kommuniziert und hierarchisch kontrolliert werden.

Rechenschaftssysteme in modernen Staaten sind immer Mischformen, bei denen lokale, professionelle und administrative Dynamiken unterschiedliches Gewicht haben können. Eine Variante ist denkbar, in der der Staat allgemeine organisatorische und regulative Rahmenrichtlinien setzt und durch Abschlussprüfungen für Schüler und Schülerinnen das Pflichtgefühl der professionellen Akteure in den Schulen weckt, alles zu tun, um möglichst vielen Schülern und Schülerinnen zum Schulerfolg zu verhelfen. Ein solches System hätte Vor- und Nachteile. Auf der einen Seite beschränkt es den notwendigen Informationsfluss zwischen den Ebenen auf passende Prüfungsaufgaben und Beurteilungskriterien und belässt die Komplexität von pädagogischen Entscheidungen in den Händen derjenigen, die über lokales und professionelles Wissen verfügen. Auf der anderen Seite lässt sich eine verwurzelte

Leistungskultur, die sich im Zusammenspiel von Schule und Familie etabliert hat, vermutlich nur schwer in diesem System dynamisieren, und mangelnde Schulleistungen, z.B. von benachteiligten sozialen Statusgruppen, bleiben tolerierbar, solange sie in konsensuale gesellschaftliche oder professionelle Erwartungen eingebettet sind.

In weiten Teilen der USA war eine Mischung aus starken lokalen Kontrollmechanismen, geringer zentraler Kontrolle der Rahmenbedingungen und schwacher Profession bis in die 60er-Jahre hinein vorherrschend (vgl. Wirt/Kirst 1997). Mit schwindender Bedeutung der kommunalen Schulbehörden und einer anhaltenden Schwäche der Profession hat sich die administrative Zentrale, sei es auf Staats- oder auf Bundesebene, in den Vordergrund geschoben. Sie macht die ungenügende Anstrengung der Profession, die dysfunktionale Lokalpolitik in den Kommunen und die niedrige Leistungserwartung für benachteiligte Schüler und Schülerinnen⁴ für Leistungsdefizite verantwortlich. Damit wird die zentrale Steuerung mit administrativen Mitteln zum Primat, und ihre Durchschlagskraft muss die mangelnde Einbeziehung der lokalen und professionellen Kräfte kompensieren.

Es gibt Indizien in der Forschung, dass die Durchschlagskraft dieses Systems nicht unbedeutend ist. Das Schulsystem ist gewissermaßen durch eine Revolution von oben in Bewegung geraten. Die Veränderungen in den staatlichen Tests zeigen an, dass das Rechenschaftssystem ernst genommen wird und die Lehrerarbeit beeinflusst. Trotz eines verbreiteten Skeptizismus machen Schulen große Anstrengungen, ihre Wachstumsziele zu erreichen. In den positiven Fallstudien wird deutlich, dass die Leistungsdefizite vormals vernachlässigter Schülergruppen sichtbar werden und ernst genommen werden. Schulen und Schulverwaltungen suchen nach Unterrichtsprogrammen, die empirisch evaluiert sind und die Testresultate verbessern können. Ein neues Management-Modell – Schulen in flexibler halb privater Trägerschaft – wird erprobt. Das Schulsystem ist also in Bewegung geraten.

Aber wie wir dokumentiert haben, hat diese Bewegung auch gravierende nachteilige Nebeneffekte. Da die verschiedenen Steuerungsinstrumente tief in die Komplexität pädagogischen Handelns vor Ort eingreifen, haben Konstruktionsfehler des Systems unabsehbare und weitreichende Folgen. Wenn die Wachstumsziele falsch gewählt werden, hat das Konsequenzen für die Anreizfunktion. Wenn Ziele gemessen an den vorhandenen Kapazitäten zu hoch angesetzt werden, stumpft die Motivation nach einer Welle der Erregung ab oder schlägt in Demoralisierung um. Dieses Szenarium scheint sich in vielen Schulen im unteren Leistungsspektrum gegenwärtig abzuspielen. Mit immer höheren Zahlen von scheiternden Schulen, gekoppelt an die zentrale Forderung nach zunehmend radikalen Lösungen, kann von einer rationalen Allokation von Ressourcen und Reformenergien für die am meisten bedürftigen

4 „The soft bigotry of low expectations“; vgl. URL: <http://www.nrrf.org/paige3-12-03.htm>; Zugriffdatum: 13.02.2012.

Schulen kaum noch die Rede sein. Mit explodierenden Zahlen von korrekturbedürftigen Schulen wird das System dysfunktional.

Gegenwärtig kann das System mit nur sehr wenigen Indikatoren operieren (vor allem mit demografischen Daten, Anwesenheit, Testdaten und Lehrertifikaten), für die die zentralen Informationssysteme Statistiken bereitstellen. Diese Indikatoren erlauben es nicht, Manipulationen der Schulen, vor allem die Exklusion „auf kaltem Wege“, zu erkennen und entsprechend zu sanktionieren. Es ist völlig unklar, was mit schwer standardisierbaren Schülergruppen geschehen soll. Werden sie vom Sanktionsregime ausgeschlossen, weil sie nicht in die simple Metrik hineinpassen, so werden Schulen, die unter Druck stehen, sie beiseiteschieben. Werden diese Schülergruppen einbezogen, hat das negative Folgen für Anreizfunktion und Motivation. Hoher Druck wirkt hier gänzlich kontraproduktiv.

Doch das entscheidende Problem des Systems ist, dass es eine Lernkultur hervorbringt, die mit gutem Unterricht wenig zu tun hat. Zumindest im unteren Leistungsspektrum des Schulsystems vollziehen sich offensichtlich eine strikte Ausrichtung an den bundesstaatlichen standardisierten *High-Stakes*-Tests und ein Wettlauf in den kognitiven Basisbereichen; beides zahlt sich nicht in entsprechenden Steigerungen in *Low-Stakes*-Tests (NAEP) aus. Der ‚Goldstandard‘ der Systemleistung ist aber der *Low-Stakes*-Test, und dieser zeigt an, dass das System sein wichtigstes Ziel, nämlich das Schließen der Leistungslücke zwischen benachteiligten und privilegierten Schülern und Schülerinnen, das „*closing of the achievement gap*“, verfehlt hat. Das US-amerikanische System der zentralen Steuerung mit administrativen Mitteln hat sich offensichtlich übernommen.

Wie könnte es weitergehen? Eine Variante wäre, das administrative Steuerungssystem besser mit einer professionellen Dynamik von Schulreform zu verknüpfen (vgl. O’Day 2002) oder lokale Stimmen stärker einzubeziehen. Das Problem ist, dass dann Interessengruppen Einfluss gewinnen, die im momentanen Diskurs als Problemträger identifiziert sind. Die Obama-Administration hat aus den Dysfunktionalitäten des NCLB-Regelwerks andere Konsequenzen gezogen. Die Funktionslogik des Primats zentraler administrativer Steuerung von Schule und Unterricht wird beibehalten; aber in den letzten zwei Jahren sind Novellierungen in allen wichtigen Funktionsbereichen geplant oder schon unternommen worden,⁵ die den Schwerpunkt des Systems verlagern. Mit der Aufgabe fixer quantitativer Zielsetzungen wird das System flexibilisiert; mit der Neuorientierung auf komplexere Standards im Bereich „*college readiness*“ wird der kognitive Horizont des Systems erweitert.⁶ Entscheidend ist, dass nun

5 Quellen: URL: <http://www.ed.gov/esea/flexibility>; URL: <http://www2.ed.gov/programs/racetothetop/>; URL: <http://www.corestandards.org/>; URL: <http://www2.ed.gov/policy/elsec/leg/blueprint/>; URL: <http://www.districtadministration.com/news/harkins>; Zugriffsdatum: 13.02.2012.

6 46 Bundesstaaten haben bis dato die neuen Standards akzeptiert. Vgl. URL: <http://www.corestandards.org/>; Zugriffsdatum: 13.02.2012.

die Hauptlast der Rechenschaft von der Organisation (Behörde, Schule) auf die einzelnen Lehrkräfte verlagert wird. Diese, so fordert die Bundesregierung von den Staaten, sollen jetzt regelmäßig mit Hilfe von standardisierten Testresultaten und standardisierten Lehrproben evaluiert werden. Evaluationen müssen in die Laufbahnentscheidungen der Behörden mit eingehen. Überdies werden Gelder für Interventionen auf die unteren fünf Prozent der Schulen beschränkt. Allerdings sind die kommunalen Schulbehörden verpflichtet, bei der Annahme von Bundesgeldern die Schulleitungen oder mindestens die Hälfte eines jeweiligen Kollegiums auszutauschen oder die Schulen in halb öffentliche *Charter Schools* umzuwandeln.

Dies ist der Plan der Regierung. Aber es ist nicht klar, ob es zu diesen Veränderungen kommen wird, da der Kongress, in dem die Skepsis an der Mikro-Steuerung des Schulwesens durch den Bund stark zugenommen hat, eine sehr viel föderalere Version der NCLB-Novellierung verabschieden könnte. Es ist klar, die Hochphase der zentralen Steuerung ist beendet: Ein Schließen der Leistungslücke wird es mit Leistungsmessung, Wachstumsquoten, Sanktionen und zentral verordneten Lösungen in den USA nicht geben. Aber die Phase einer zentralen Kontrolle der professionellen Arbeit hat vermutlich gerade erst begonnen. Und eingedenk der strukturellen und politischen Voraussetzungen wird das Land seine Vorreiterrolle in dieser Dimension so schnell nicht verlieren.

Was kann ein deutsches Lesepublikum aus den amerikanischen Erfahrungen lernen? Die USA beenden gerade eine Etappe in einem zwanzigjährigen Großversuch, der mit vier Grundannahmen begann, die wir provisorisch *Willens-Hypothese*, *Vereinfachungs-Hypothese*, *Dynamisierungs-Hypothese* und *Systemische-Reform-Hypothese* nennen wollen.

Die *Willens-Hypothese* besagt, dass der mangelnde Wille der Profession für die Leistungsdefizite des Systems verantwortlich ist und dass Anreize das zentrale Mittel sind, um das System zu verbessern. Wir wissen nun, dass diese Hypothese so nicht stimmt. Das System hat es zwar geschafft, Schulen mit Entschlossenheit auf die staatlichen Tests auszurichten und die Testresultate zu verbessern, aber vergleichbar substantielle Lernleistungen von Schülern und Schülerinnen, die sich in *Low-Stakes*-Tests zeigen müssten, scheint dieses Anreizsystem nicht hervorzubringen.

Die *Vereinfachungs-Hypothese* wurde von Management- und Psychometrie-Experten und -Expertinnen formuliert. Sie besagt, dass wir durchaus ein komplexes System menschlichen Miteinanders, nämlich den Unterricht, mit Hilfe einer begrenzten Anzahl wirkmächtiger quantitativer Indikatoren humanisieren können. Das amerikanische Großexperiment hat uns gelehrt, wie schwer es ist, die exakten numerischen Größen mit den diffuseren professionellen Werten menschlicher Fürsorge, leidenschaftlicher Wissensvermittlung und gutem Unterricht in Verbindung zu bringen. Wir haben Fallstudien erwähnt, in denen das Rechenschaftssystem machtvoll auf die-

se nicht quantitativ messbaren Werte professionellen Arbeitens ausstrahlt, aber die Standardreaktion, so scheint es, besteht darin, den kürzesten Weg zwischen geforderten Messeinheiten und antrainiertem Lehrer- und Schülerverhalten zu wählen; damit ist der Wettlauf nach unten eröffnet.

Die *Dynamisierungs-Hypothese* besagt, dass eine Revolution von oben benötigt wird, um die Starrheit der mittleren und unteren Ebenen des Systems zu verflüssigen. Das amerikanische Großexperiment hat gezeigt, dass diese Hypothese insofern stimmt, als Initiativen der Zentrale in der Tat untere Ebenen in große Turbulenzen stürzen können. Aber die Treffsicherheit dieses Impulses ist wenig überzeugend, wenn die Macht von Sanktionen auf eine wenig entwickelte Reforminfrastruktur trifft.

Die letzte Hypothese, die hier angesprochen werden soll, ist die *Systemische-Reform-Hypothese*. Sie postuliert, dass Systemelemente, wie Lehrpläne, Tests, Unterrichtsmaterialien und Lehrerfortbildungen aufeinander abgestimmt sein sollten, um die Entwicklungsanstrengungen der Lehrenden in einer Struktur von Vorhersehbarkeit, Datenevidenz und externer Unterstützung abzusichern. Wir haben Anhaltspunkte dafür, dass sich diese Hypothese, nach Lage der Daten in den USA, als haltbar erweisen könnte.

Literatur

- Abedi, J. (2004): The No Child Left Behind Act and English Language Learners: Assessment and Accountability Issues. In: *Educational Researcher* 33, H. 1, S. 4-14.
- Abrams, L./Pedulla, J./Madaus, G. (2003): Views from the Classroom: Teachers' Opinions of Statewide Testing Programs. In: *Theory into Practice* 42, H. 1.
- Amrein, A.L./Berliner, D.C. (2002): High-Stakes Testing, Uncertainty, and Student Learning. In: *Education Policy Analysis Archives* 10, H. 18. URL: <http://epaa.asu.edu/epaa/v10n18/>; Zugriffsdatum: 22.08.2003.
- Anagnostopoulos, D. (2003): The New Accountability, Student Failure, and Teachers' Work in Urban High Schools. In: *Educational Policy* 17, H. 3, S. 291-316.
- Asimov, N. (2008, September 5): State Falling Way behind No Child Left Behind. *San Francisco Chronicle*, S. A-1.
- Au, W. (2007): High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. In: *Educational Researcher* 36, H. 5, S. 258-267.
- Balfanz, R./Legters, N.E. (2004): Locating the Dropout Crisis: Which High Schools Produce the Nation's Dropouts? In: Orfield, G. (Hrsg.): *Dropouts in America: Confronting the Graduation Rate Crisis*. Cambridge, MA: Harvard Education Press, S. 57-84.
- Balfanz, R./Legters, N.E. (2008): NCLB and Reforming the Nation's Lowest-Performing High Schools. In: Sunderman, G.L. (Hrsg.): *Holding NCLB Accountable: Achieving Accountability, Equity, and School Reform*. Thousand Oaks, CA: Corwin Press, S. 191-207.
- Batt, L./Kim, J./Sunderman, G.L. (2005): Limited English Proficiency Students: Increased Accountability under NCLB. Cambridge, MA: The Civil Rights Project at Harvard University.

- Blanc, S. (2003, Summer): Supports and Challenges in the Multiple Provider Model: Principals Offer Mixed Perspectives on First Year of Takeover. Philadelphia Public School Notebook.
- Booher-Jennings, J. (2005): Below the Bubble: "Educational Triage" and the Texas Accountability System. In: *American Educational Research Journal* 42, H. 2, S. 231-268.
- Bracey, G. (2002): The Market in Theory Meets the Market in Practice: The Case of Edison Schools. Education Policy Studies Laboratory: Education Policy Research Unit (EPRU), College of Education, Arizona State University. URL: <http://www.asu.edu/educ/eps/EPRL/documents/EPRL%202002-107/EPRL-0202-107-EPRU.htm>; Zugriffsdatum: 02.11.2005.
- Brady, R. (2003): *Can Failing Schools be Fixed?* Washington, DC: Thomas B. Fordham Foundation.
- Braun, H. (2004): Reconsidering the Impact of High-Stakes Testing. In: *Education Policy Analysis Archives* 12, H. 1. URL: <http://epaa.asu.edu/epaa/v12n1/>; Zugriffsdatum: 13.02.2012.
- Buchanan, B. (2006): *Turnover at the Top: Superintendent Vacancies and the Urban School.* Lanham, PA: Rowman & Littlefield.
- Callahan, R. (1962): *Education and the Cult of Efficiency.* Chicago: University of Chicago Press.
- Carnoy, M./Loeb, S. (2002): Does External Accountability Affect Student Outcomes? In: *Education Evaluation and Policy Analysis* 24, S. 305-332.
- Center for Education Policy (2007): *Moving beyond Identification: Assisting Schools in Improvement.* Washington, DC: Author.
- Center for Education Policy (2008a): *Has Student Achievement Increased since 2002? State Test Score Trends through 2006-07.* Washington, DC: Author.
- Center for Education Policy (2008b): *Managing More than a Thousand Remodeling Projects: School Restructuring in California.* Washington, DC: Author.
- Cohen, D.K./Spillane J. (1992): Policy and Practice: The Relations Between Governance and Instruction. In: *Review of Research in Education* 18, S. 3-49.
- Coltrane, B. (2002, November): English Language Learners and High-Stakes Tests: An Overview of the Issues. URL: <http://www.cal.org/ericcl/DIGEST>; Zugriffsdatum: 13.05.2004.
- Conley, D./Goldman, P. (2000): Half Full, Half Empty? Educator Responses over Time to State-level Systematic Reforms. In: *International Journal of Educational Reform* 9, H. 3, S. 249-260.
- Cuban, L. (1990): Reforming Again, Again, and Again. In: *Educational Researcher* 19, H. 1, S. 3-13.
- Darling-Hammond, L. (1990): Instructional Policy and Practice: The Power of the Bottom over the Top. In: *Educational Evaluation and Policy Analysis* 12, H. 3, S. 339-347.
- Darling-Hammond, L. (2004): Standards, Accountability, and School Reform. In: *Teachers College Record* 106, H. 6, S. 1047-1085.
- Debray, E. (2006): *Politics, Ideology, and Education: Federal Policy during the Clinton and Bush Administrations.* New York: Teachers College Press.
- Diamond, J.B./Spillane, J. (2004): High-Stakes Accountability in Urban Elementary Schools: Challenging or Reproducing Inequality? In: *Teachers College Record* 106, S. 1145-1176.
- Elmore, R.F. (2004): *School Reform from the Inside Out: Policy, Practice and Performance.* Cambridge, MA: Harvard Education Press.

- Finnigan, K./Gross, B. (2007): Do Accountability Policy Sanctions Influence Teacher Motivation? Lessons from Chicago's Low-Performing School. In: *American Educational Research Journal* 44, H. 3, S. 594-629.
- Firestone, W./Monfils, L./Schorr, R. (2004): *The Ambiguity of Teaching to the Test: Standards, Assessment, and Educational Reform*. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Fuhrman, S./O'Day, J. (Hrsg.) (1996): *Rewards and Reform*. San Francisco: Jossey-Bass.
- Fullan, M. (2003): *Change Forces with a Vengeance*. London/New York: Routledge Farmer.
- Fuller, B./Gesicki, K./Kang, E./Wright, J. (2006): Is the No Child Left Behind Act Working? The Reliability of how States Track Achievement. Working Paper 06-1 Policy Analysis for California Education, PACE. 3653 Tolman Hall, Berkeley, CA 94720-1670. Tel: 510-642-7223; Fax: 510-642-9148; E-Mail: pace@berkeley.edu. URL: http://pace.berkeley.edu/pace_publications.html.
- Fuller, B./Wright, J./Gesicki, K./Kang, E. (2007): Gauging Growth: How to Judge No Child Left Behind? In: *Educational Researcher* 36, H. 5, S. 268-278.
- GAO (U.S. Government Accountability Office) (2007): *No Child Left Behind Act: Education Should Clarify Guidance and Address Potential Compliance Issues for Schools in Corrective Action and Restructuring Status*. Washington, DC: Author.
- Gill, B./Zimmer, R./Christman, J./Blanc, S. (2007): *State Takeover, School Restructuring, Private Management, and Student Achievement in Philadelphia*. Santa Monica: Rand.
- Haney, W.M. (2000): The Myth of the Texas Miracle in Education. In: *Education Policy Analysis Archives* 8, H. 41. URL: <http://epaa.asu.edu/epaa/v8n41/>; Zugriffsdatum: 22.08.2006.
- Haney, W.M. (2008): Evidence on Education under NCLB (and How Florida Boosted NAEP Scores and Reduced the Race Gap). In: Sunderman, G.L. (Hrsg.): *Holding NCLB Accountable: Achieving Accountability, Equity, and School Reform*. Thousand Oaks, CA: Corwin Press, S. 91-102.
- Herman, J.L. (2004): The Effects of Testing on Instruction. In: Fuhrman, S./Elmore, R. (Hrsg.): *Redesigning Accountability*. New York: Teachers College Press, S. 141-146.
- Hess, G.A. (2003): Reconstitution – Three Years Later: Monitoring the Effect of Sanctions on Chicago High Schools. In: *Education and Urban Society* 35, H. 3, S. 300-327.
- Hill, P./Lake, R./Celio, M. (2002): *Charter Schools and Accountability in Public Education*. Washington, DC: Brookings Institution Press.
- Johnson, S. (2007): *Finders and Keepers: Helping New Teachers Survive and Thrive in Our Schools*. Indianapolis, IN: Jossey-Bass.
- Kelley, C./Protsik, J. (1997): Risk and Reward: Perspectives on the Implementation of Kentucky's School-based Performance Award Program. In: *Educational Administration Quarterly* 33, H. 4, S. 474-505.
- Kieffer, M.J./Lesaux, N.K./Snow, C.E. (2008): Promises and Pitfalls: Implications of NCLB for Identifying, Assessing, and Educating English Language Learners. In: Sunderman, G.L. (Hrsg.): *Holding NCLB Accountable: Achieving Accountability, Equity, and School Reform*. Thousand Oaks, CA: Corwin Press, S. 57-74.
- Kim, J./Sunderman, G.L. (2004): *Does NCLB Provide Good Choices for Students in Low Performing Schools?* Cambridge, MA: The Civil Rights Project at Harvard University.
- Kim, J./Sunderman, G.L. (2005): Measuring Academic Performance under NCLB: Implications for Educational Equity. In: *Educational Researcher* 34, H. 8, S. 3-13.
- Koretz, D.M. (2008): The Pending Reauthorization of NCLB: An Opportunity to Rethink the Basic Strategy. In: Sunderman, G.L. (Hrsg.): *Holding NCLB Accountable: Achieving Accountability, Equity, and School Reform*. Thousand Oaks, CA: Corwin Press, S. 9-26.

- Koretz, D.M./Barron, S.I./Mitchell, K.J./Stecher, B.M. (1996): Perceived Effects of the Kentucky Instructional Results Information System. Santa Monica, CA: RAND.
- Lee, J. (2006): Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps: An In-depth Look into National and State Reading and Math Outcome Trends. Cambridge, MA: The Civil Rights Project at Harvard University.
- Lee, J. (2007): The Testing Gap: Scientific Trials of Test-driven School Accountability Systems for Excellence and Equity. Charlotte, NC: Information Age.
- Lee, J. (2008): Is Test-driven External Accountability Effective? Synthesizing the Evidence from Cross-State Causal-Comparative and Correlational Studies. In: Review of Educational Research 78, S. 608-644.
- Leithwood, K./Steinbach, R./Jantzi, D. (2002): School Leadership and Teachers' Motivation To Implement Accountability Policies. In: Educational Administration Quarterly 38, H. 1, S. 94-119.
- Linn, R.L. (2005): Conflicting Demands of No Child Left Behind and State Systems: Mixed Messages about School Performance. In: Education Policy Analysis Archives 13, H. 33.
- Loeb, H./Knapp, M./Elfers, A. (2008): Teachers' Response to Standards-based Reform: Probing Reform Assumptions in Washington State. In: Education Policy Analysis Archives 16, H. 8, S. 1-29.
- Louis, K.S./Febey, K./Schroeder, R. (2005): State-Mandated Accountability in High Schools: Teachers' Interpretations of a New Era. In: Educational Evaluation and Policy Analysis 27, S. 177-204.
- Malen, B./Croninger, R./Muncey, D./Redmond-Jones, D. (2002): Reconstituting Schools: „Testing“ the „Theory of Action“. In: Educational Evaluation and Policy Analysis 24, H. 2, S. 113-132.
- Marion, S.T./White, C./Carlson, D./Erpenbach, W.J./Rabinowitz, A./Sheinker, J. (2002): Making Valid and Reliable Decisions in Determining Adequate Yearly Progress. Washington, DC: Council of Chief State School Officers.
- Mathis, W.J. (2009): NCLB's Ultimate Restructuring Alternatives: Do They Improve the Quality of Education? East Lansing, MI: The Great Lakes Center for Education Research & Practice.
- McDermott, K.A. (2004): Incentives, Capacity, and Implementation: Evidence from Massachusetts Education Reform. In: Journal of Administration Research and Theory 16, S. 45-65.
- McLaughlin, M. (1990): The Rand Change Agent Study Revisited. In: Educational Researcher 19, H. 9, S. 11-16.
- McNeil, L. (2000): Contradictions of School Reform. New York/London: Routledge.
- Mintrop, H. (2004): Schools on Probation: How Accountability Works (and Doesn't Work). New York: Teachers College.
- Mintrop, H. (2008): Low-Performing Schools' Programs and State Capacity Requirements: Meeting the NCLB Educational Goals. In: Sunderman, G.L. (Hrsg.): Holding NCLB Accountable: Achieving Accountability, Equity, & School Reform. Thousand Oaks, CA: Corwin Press, S. 137-151.
- Mintrop, H./Sunderman, G. (2009): Predictable Failure of Sanctions-driven Accountability for School Improvement – and Why We May Retain It Anyway. In: Educational Researcher 38, H. 5, S. 353-364.
- Mintrop, H./Trujillo, T. (2005): Corrective Action in Low Performing Schools: Lessons for NCLB Implementation from First-Generation Accountability Systems. In: Education Policy Analysis Archives 13, H. 48.
- Mintrop, H./Trujillo, T. (2007): The Practical Relevance of Accountability Systems for School Improvement: A Descriptive Analysis of California Schools. In: Educational Evaluation and Policy Analysis 29, H. 4, S. 319-352.

- Neal, D./Schanzenbach, D.W. (2008): *Left Behind by Design: Proficiency Counts and Test-based Accountability*. Chicago: University of Chicago.
- Neild, R.C./Spiridakis, K. (2002): *Preliminary Results: Philadelphia New Teacher Survey*. Philadelphia, PA: Research for Action.
- Neild, R.C./Useem, B./Travers, E./Lesnick, J. (2003): *Once & for All: Placing a Highly Qualified Teacher in Every Philadelphia Classroom*. Philadelphia, PA: Research for Action.
- Nichols, S./Glass, G./Berliner, D. (2006): *High-Stakes Testing and Student Achievement: Does Accountability Pressure Increase Student Learning?* In: *Educational Policy Analysis Archives* 17, H. 1, S. 1-72.
- Novak, J.F.B. (2003, December): *Penalizing Diverse Schools? Similar Test Scores but Different Students Bring Federal Sanctions*. Berkeley: Policy Analysis for California Education.
- NYSED (New York State Education Department), Office of New York City School Improvement and Community Services (2003): *2002-03 Registration Review Initiative: A Summary of the Registration Review Report Findings*. New York City: Author.
- O'Day, J. (2002): *Complexity, Accountability, and School Improvement*. In: *Harvard Educational Review* 72, H. 3, S. 293-329.
- Oldham, G./Hackman, J.R. (2010): *Not What It Was and not What It Will Be: the Future of Job Design Research*. In: *Journal of Organizational Behavior* 31, H. 2-3, S. 463-479.
- Owens, A./Sunderman, G.L. (2006): *School Accountability Under NCLB: Aid or Obstacle for Measuring Racial Equity?* Cambridge, MA: The Civil Rights Project at Harvard University.
- Protheroe, N. (2008): *The Impact of Fidelity of Implementation in Effective Standards-based Instruction*. In: *Principal* 87, H. 6, S. 38-41.
- Raymond, M.E./Hanushek, E.A. (2003): *High-stakes Research: Accountability Works after All*. In: *Education Next* 3, H. 3, S. 48-55.
- Saltman, K.J. (2005): *The Edison Schools*. New York: Routledge.
- Scheerens, J./Bosker, R. (1997): *The Foundations of Educational Effectiveness*. Oxford, GB: Pergamon.
- Sennett, R. (2006): *The Culture of New Capitalism*. New Haven: Yale University Press.
- Skrla, L./Scheurich, J. (2004): *Displacing Deficit Thinking in School District Leadership*. In: Skrla, L./Scheurich, J. (Hrsg.): *Educational Equity and Accountability*. New York: Routledge Farmer, S. 109-132.
- Slavin, R. (2005): *Evidence-based Reform: Advancing the Education of Students at Risk*. Report Prepared for: *Renewing Our Schools, Securing Our Future; a National Task Force on Education*. o.V.
- Sunderman, G.L. (2001): *Accountability Mandates and the Implementation of Title I Schoolwide Programs: A Comparison of Three Urban Districts*. In: *Educational Administration Quarterly* 37, S. 503-532.
- Sunderman, G.L./Kim, J.S./Orfield, G. (2005): *NCLB Meets School Realities: Lessons from the Field*. Thousand Oaks, CA: Corwin Press.
- Sunderman, G.L./Orfield, G. (2006): *Domesticating a Revolution: No Child Left Behind and State Administrative Response*. In: *Harvard Educational Review* 76, H. 4, S. 526-556.
- Sunderman, G.L./Tracey, C.A./Kim, J./Orfield, G. (2004): *Listening to Teachers: Classroom Realities and No Child Left Behind*. Cambridge, MA: The Civil Rights Project at Harvard University.
- The Brown Center (2003): *Charter Schools: Achievement, Accountability, and the Role of Expertise*. Washington, DC: Brookings Institution.
- The Education Trust (2004): *Measured Progress: Achievement Rises and Gaps Narrow, but Slowly*. Washington, DC: Author.

- The Education Trust (2005): Stalled in Secondary: A Look at Student Achievement since the No Child Left Behind Act. Washington, DC: Author.
- Timar, T. (1997): The Institutional Role of State Education Departments. In: American Journal of Education 105, H. 3, S. 231-260.
- Timar, T./Kirp, D. (1988): State Efforts to Reform Schools: Treading Between a Regulatory Swamp and an English Garden. In: Educational Evaluation and Policy Analysis 10, H. 2, S. 75-88.
- Travers, E. (2003a): Characteristics of Schools under Diverse Providers 2002-2003. Philadelphia, PA: Research for Action.
- Travers, E. (2003b): Philadelphia School Reform: Historical Roots and Reflections on the 2002-2003 School Year under State Takeover: A Talk Initially Presented at The Friends Association of Higher Education and Friends Council on Education Conference at Swarthmore College and Pendle Hill. Penn GSE Perspectives on Urban Education 2, H. 1. URL: <http://www.urbanedjournal.org/commentaries/c0007.pdf>; Zugriffsdatum: 02.11.2005.
- Tyack, D. (1974): The One Best System: A History of Urban Education. Cambridge, MA: Harvard University Press.
- Useem, E. (2005): Learning from Philadelphia's School Reform: What Do the Research Findings Show so Far? Philadelphia, PA: Research for Action.
- Valenzuela, A. (2005): Leaving Children Behind: How "Texas-style" Accountability Fails Latino Youth. Albany: State University of New York Press.
- Verstegen, D./Knoeppel, R. (1998): Equal Education under the Law: School Finance Reform and the Courts. In: Journal of Law and Politics 14, H. 3, S. 555-592.
- Weiler, H. (1990): Comparative Perspectives in Decentralization in Education: An Exercise in Contradiction? In: Educational Evaluation and Policy Analysis 12, H. 4, S. 433-448.
- Wirt, F./Kirst, M. (1997): The Political Dynamics of American Education. Berkeley, CA: McCutchan.

Heinrich Mintrop, Associate Professor, Educational Policy and Leadership, Director of the Doctoral Program in Leadership for Educational Equity, Graduate School of Education, University of California, Berkeley, USA.

Anschrift: 3647 Tolman Hall, University of California, Berkeley, CA. 94720, USA
E-Mail: mintrop@berkeley.edu

Gail Sunderman, Ph.D., Senior Research Scientist at the George Washington University Center for Equity and Excellence in Education, Director of the Mid-Atlantic Equity Center.

Anschrift: MAEC at The George Washington University Center for Equity and Excellence in Education, 1555 Wilson Blvd., Suite 515, Arlington, Virginia 22209, USA
E-Mail: glsunderman@yahoo.com